

# Topological determinants of complex networks spectral properties: structural and dynamical effects

Claudio Castellano<sup>1</sup> and Romualdo Pastor-Satorras<sup>2,\*</sup>

<sup>1</sup>*Istituto dei Sistemi Complessi (ISC-CNR), Via dei Taurini 19, I-00185 Roma, Italy*

<sup>2</sup>*Departament de Física, Universitat Politècnica de Catalunya, Campus Nord B4, 08034 Barcelona, Spain*

The largest eigenvalue of a network's adjacency matrix and its associated principal eigenvector are key elements for determining the topological structure and the properties of dynamical processes mediated by it. We present a physically grounded expression relating the value of the largest eigenvalue of *any network* to the largest eigenvalue of two network subgraphs, considered as isolated: The hub with its immediate neighbors and the densely connected set of nodes with maximum  $K$ -core index. We validate this formula showing that it predicts with good accuracy the largest eigenvalue of a large set of synthetic and real-world topologies, with no exception. We also present evidence of the consequences of these findings for broad classes of dynamics taking place on the networks. As a byproduct, we reveal that the spectral properties of heterogeneous networks built according to the linear preferential attachment model are qualitatively different from those of their static counterparts.

## I. INTRODUCTION

The spectral properties of complex topologies [1] play a crucial role in our understanding of the structure and function of real networked systems. Various matrices can be constructed for any given network, their spectral properties accounting for different topological or functional features. Thus, for example, the Laplacian matrix is related to diffusive and random walk dynamics on networks [2], the modularity matrix plays a role in community identification on networks [3], while the non-backtracking or Hashimoto matrix governs percolation [4]. Among all matrices associated to networks, the simplest and possibly most studied is the adjacency matrix  $A_{ij}$ , taking the value 1 whenever nodes  $i$  and  $j$  are connected, and zero otherwise. Particular interest in this case is placed on the study of the principal eigenvector  $\{f_i\}$  (PEV), defined as the eigenvector of the adjacency matrix with the largest eigenvalue  $\Lambda_M$  (LEV). This interest is twofold. On the one hand, the PEV is one of the fundamental measures of node importance or centrality [5]. The centrality of a node can be defined based on the number of other different vertices that can be reached from it, or the role it plays in connecting together different parts of the network. From a more sociological point of view, a node is central if it is connected to other central nodes. From this definition arises the notion of eigenvector centrality of a node [6], that coincides with the corresponding component of the PEV. On the other hand, the LEV plays a pivotal role in the behavior of many dynamical systems on complex networks, such as epidemic spreading [7], synchronization of weakly coupled oscillators [8], weighted percolation on directed networks [9], models of genetic control [10], or the dynamics of excitable elements [11]. In this kind of dynamical processes, the LEV is related, through different analytical techniques, to the critical point at which a transition between different phases takes place: In terms of some generic control parameter  $\lambda$ , a critical point  $\lambda_c$  is found to be in general inversely proportional to the LEV  $\Lambda_M$ .

The possibility to know the position of such transition points in terms of simple network topological properties is of great

importance, as it allows to predict the system macroscopic behavior or optimize the network to control processes on it. This has triggered an intense activity [12–15], particularly in the case of networks with heterogeneous topology, such as power-law distributed networks with a degree distribution of the form  $P(q) \sim q^{-\gamma}$  [16]. Among these efforts, in their seminal work [17], Chung, Lu and Vu (CLV) have rigorously proven, for a model with power-law degree distribution, that the LEV can be expressed in terms of the maximum degree  $q_{\max}$  present in the network and the first two moments of the degree distribution. This is a remarkable achievement, as it allows to draw predictions in the analysis of dynamics on networks, without recurring to the cumbersome numerical evaluation of the LEV.

Here we show that while the CLV theory provides in some cases an excellent approximation to the LEV, specially in the case of random uncorrelated networks, it can fail considerably in other cases. In order to provide better estimates, we reinterpret the CLV result in terms of the competition among different subgraphs in the networks. This insight leads us to the formulation of a modified form of the CLV theory, that captures the behavior of the LEV more generally, including the case of real correlated networks, and reduces to the CLV form in the case of random uncorrelated networks. We show that our generalized expression perfectly predicts the LEV for linear preferential attachment growing networks (for which the original CLV form fails) and provides an excellent approximation for the LEV of real-world networks. Finally, we show that our modified expression predicts reliably the critical point of dynamical processes on a large set of synthetic and real-world networks, with no exception.

## II. THE CHUNG-LU-VU FORMULA FOR THE LARGEST EIGENVALUE

In Ref. [17], the authors consider a class of network models with expected degree distribution. That is, starting from a predefined degree distribution  $P(q)$ , one generates expected degrees  $\tilde{q}_i$  for each node, drawn from  $P(q)$ , and creates an actual network by joining every pair of nodes  $i$  and  $j$  with probability  $\tilde{q}_i \tilde{q}_j / \sum_r \tilde{q}_r$ . The resulting network has a degree

\* Corresponding author: romualdo.pastor@upc.edu

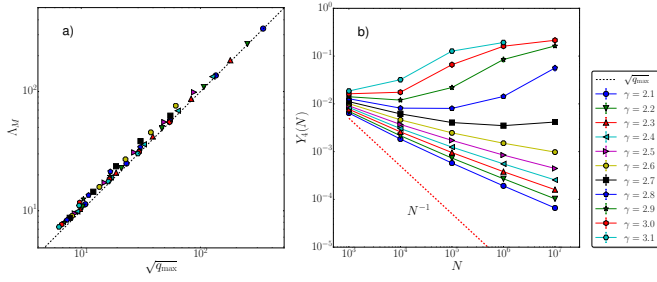


Figure 1. (a) Largest eigenvalue  $\Lambda_M$  as a function of  $\max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}$  in uncorrelated power-law UCM networks with different degree exponent  $\gamma$  and network size  $N$ . (b) Inverse participation ratio  $Y_4(N)$  as a function of  $N$  in uncorrelated power-law UCM networks with different degree exponent  $\gamma$ . Each point in both graphs corresponds to an average over 100 independent network realization. Error bars are smaller than symbol sizes. Networks have a minimum degree  $m = 3$ .

distribution with the same functional form as the imposed  $P(q)$  and lacks degree correlations, since the condition  $\tilde{q}_i^2 < \sum_r \tilde{q}_r$  is imposed in the construction [17, 18]. This algorithm is a variation of the classical configuration model [14], cast in terms of a hidden variables model [19]. For this model network, and any arbitrary degree distribution, the authors in Ref. [17] rigorously prove that the largest eigenvalue of the corresponding adjacency matrix takes the form (see also [20])

$$\Lambda_M \simeq \begin{cases} a_1 \sqrt{q_{\max}} & \text{if } \sqrt{q_{\max}} > \frac{\langle q^2 \rangle}{\langle q \rangle} \ln^2(N) \\ a_2 \frac{\langle q^2 \rangle}{\langle q \rangle} & \text{if } \frac{\langle q^2 \rangle}{\langle q \rangle} > \sqrt{q_{\max}} \ln(N) \end{cases}, \quad (1)$$

where  $N$  is the network size,  $q_{\max}$  is the maximum degree in the networks, and  $a_i$  are constants of order 1. In the case of scale-free networks, the maximum degree is a growing function of  $N$ , that for uncorrelated networks [21] takes the value  $q_{\max} \sim N^{1/2}$  for  $\gamma \leq 3$  and  $q_{\max} \sim N^{1/(\gamma-1)}$  for  $\gamma > 3$  [18]. The algebraic increase of  $q_{\max}$  allows to disregard the logarithmic terms in Eq. (1), leading to the simpler expression

$$\Lambda_M \approx \max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}, \quad (2)$$

valid for any value of  $\gamma$ . For power law distributed networks, the second moment of the degree distribution scales as  $\langle q^2 \rangle \sim q_{\max}^{3-\gamma}$  for  $\gamma \leq 3$  and  $\langle q^2 \rangle \sim \text{const.}$  for  $\gamma > 3$ . Combining this result with the expression for the maximum degree, we can write the more explicit result

$$\Lambda_M \approx \begin{cases} \sqrt{q_{\max}} & \text{if } \gamma > 5/2 \\ \frac{\langle q^2 \rangle}{\langle q \rangle} & \text{if } \gamma < 5/2 \end{cases}. \quad (3)$$

The Chung *et al.* (CLV) prediction for the largest eigenvalue of a power-law distributed network, Eq. (2), turns out to be very accurate for random uncorrelated static networks. Indeed, in Fig. 1a we present a scatter plot of  $\Lambda_M$  computed using the power iteration method (UCM) [22] in random uncorrelated power-law networks generated with the uncorrelated configuration model [23], for different values of the degree exponent  $\gamma$

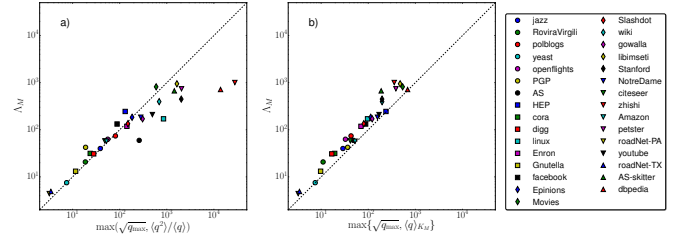


Figure 2. (a) Largest eigenvalue  $\Lambda_M$  as a function of  $\max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}$  computed for 31 different real-world networks. (b) Largest eigenvalue  $\Lambda_M$  as a function of  $\max\{\sqrt{q_{\max}}, \langle q \rangle_{K_M}\}$  computed for the same real-world networks. Networks are ordered by increasing network size.

and network size  $N$ , as a function of the numerically estimated value of  $\max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}$ . The fit to the CLV theory is almost perfect, with only very small deviations for the smallest network sizes.

In order to test the generality of this prediction, we have also considered a large set of real-world networks, of widely different origin, size and topological features (see Appendix F). In Fig 2a we plot the LEV of these networks as a function of the numerically estimated quantity  $\max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}$ . The result is quite clear: While in some cases the CLV prediction works quite well, in others it provides a large overestimation of the actual value of the LEV, of about one order of magnitude or larger. This discrepancy is particularly strong in the case of the zhishi, dbpedia, Stanford, linux, and AS networks.

### III. GENERALIZED FORMULA FOR THE LARGEST EIGENVALUE

We can understand the origin of the violations of the CLV formula observed Fig. 2a and provide an improved version, by reconsidering the observations made in Refs. [24, 25]. In these works it is shown that the two types of scaling of the LEV in the CLV formula for uncorrelated networks (either proportional to  $\langle q^2 \rangle / \langle q \rangle$  or to  $\sqrt{q_{\max}}$ ) are the manifestation of the two alternative ways in which the PEV can become localized in the network (see Appendix A). For  $\gamma > 5/2$  the PEV is localized around the node with largest degree in the network (the hub) and the scaling of  $\Lambda_M$  is given by  $\sqrt{q_{\max}}$ ; for  $\gamma < 5/2$ , on the other hand, the PEV becomes localized (in the sense discussed in Ref. [25]) on the core of nodes of maximum index  $K_M$  in the  $K$ -core decomposition of the network [26, 27] (see Appendix B); the associated LEV is then given by  $\langle q^2 \rangle / \langle q \rangle$ . This picture is confirmed in Fig. 1b, where we study the localization of the PEV, of components  $\{f_i\}$  assumed to be normalized as  $\sum_i f_i^2 = 1$ . The analysis is performed by plotting the inverse participation ratio [25, 28, 29]  $Y_4(N)$  as a function of the network size (see Appendix A). As we can check, for  $\gamma > 5/2$ ,  $Y_4(N)$  goes to a constant for  $N \rightarrow \infty$ , indicating localization in a finite set of nodes, that coincide with the hub. On the other hand, for  $\gamma < 5/2$  the inverse participation ratio decreases algebraically with network size, with an exponent  $\alpha$  smaller than 1, indicating localization in a

sub-extensive set of nodes, which coincide with the maximum  $K$ -core[25].

This observation can be interpreted in the following terms: The actual value of the LEV in the whole network is the result of the competition among two different subgraphs. The node with largest degree  $q_{\max}$  (hub), together with its immediate neighbors, form a star graph which, in isolation, has a largest eigenvalue given by  $\Lambda_M^{(h)} = \sqrt{q_{\max}}$ . On the other hand, the maximum  $K$ -core, of degree  $K_M$  is a densely interconnected, essentially degree-homogeneous subgraph [24]. As such, its largest eigenvalue is given by its internal average degree,  $\Lambda_M^{(K_M)} = \langle q \rangle_{K_M}$ . In the case of uncorrelated networks, this average degree is given by  $\langle q^2 \rangle / \langle q \rangle$  [27]. These two subgraphs, and their respective largest eigenvalues,  $\Lambda_M^{(h)}$  and  $\Lambda_M^{(K_M)}$ , compete in order to set the scaling of the LEV of the whole network: the global LEV coincides with the subgraph LEV that is larger.

We hypothesize that for any network, also for correlated ones, the same competition sets the overall LEV value. The largest eigenvalue of the star graph centered around the hub is trivially still equal to  $\sqrt{q_{\max}}$ . What changes in general topologies, and in particular in correlated networks, is the expression of the largest eigenvalue associated to the maximum  $K$ -core. One can realistically assume that the maximum  $K$ -core is in general degree-homogeneous (see the heterogeneity parameter of the maximum  $K$ -core or real-world networks in Table I, which is always much smaller than 1). What cannot be taken for granted in general is the identification between  $\langle q \rangle_{K_M}$  and  $\langle q^2 \rangle / \langle q \rangle$ . We thus conjecture that the LEV in generic networks can be expressed as

$$\Lambda_M \approx \max(\sqrt{q_{\max}}, \langle q \rangle_{K_M}). \quad (4)$$

Eq. (4) is the central result of our paper. Notice that Eq. (4) includes Eq. (2) as the particular case when  $\langle q \rangle_{K_M} = \langle q^2 \rangle / \langle q \rangle$ , which is true in uncorrelated networks.

In Fig. 2b we check the validity of the proposed generalized scaling for the LEV in the case of the 31 real-world networks considered above. Comparing with Fig. 2a, the generalized formula provides a much better overall fitting to the real value of the LEV than the original CLV expression, and therefore represents a better prediction for the behavior of this quantity. The overall improvement of our prediction versus the original CLV one can be established by comparing the respective Pearson correlation coefficients of the predictions and the real observed measurements. We obtain a Pearson coefficient  $r = 0.893$  for our improved prediction Eq. (4), while the original CLV, Eq. (2) takes  $r = 0.617$ .

#### IV. THE CASE OF LINEAR PREFERENTIAL ATTACHMENT NETWORKS

Growing network models provide a particularly interesting testbed for the conjecture presented above. We focus in particular on linear preferential attachment (LPA) networks [16, 30], generated starting from a fully connected nucleus of  $m + 1$  nodes and adding at every time step a new node with  $m$  new

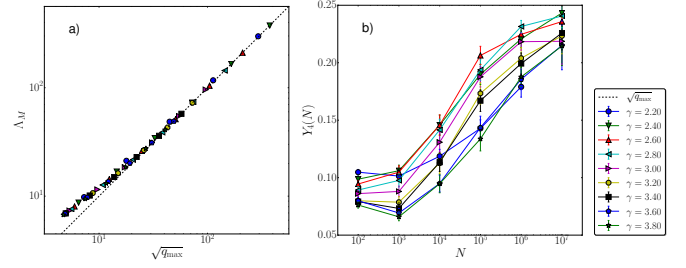


Figure 3. (a) Largest eigenvalue  $\Lambda_M$  as a function of  $\sqrt{q_{\max}}$  in LPA networks with different degree exponent  $\gamma$ . (b) Inverse participation ratio  $Y_4(N)$  as a function of  $N$  in LPA networks with different degree exponent  $\gamma$ . Each point in both graphs corresponds to an average over 100 independent network realization. Error bars are smaller than symbol sizes.

edges connected to  $m$  old nodes. For the vertex introduced at time  $t$ , each of its emanating edges is connected to an existing vertex  $s$ , introduced at time  $s < t$ , with probability

$$\Pi_s(t) = \frac{q_s(t) + a}{\sum_j [q_j(t) + a]}, \quad (5)$$

where  $q_s(t)$  is the degree, measured at time  $t$ , of the node introduced at time  $s$ . The factor  $a$  takes into account the possible initial attractiveness of each node, prior to receiving any connection. Large LPA networks are characterized by a power-law degree distribution  $P(k) \sim k^{-\gamma}$ , with a degree exponent  $\gamma = 3 + \frac{a}{m}$  [31] and average degree  $\langle q \rangle = 2m$ . It is thus possible to tune the degree exponent in the range  $2 < \gamma < \infty$  by changing the attractiveness parameter in the range  $-m < a < \infty$ . The power-law form extends up to the maximum degree  $q_{\max}$  that depends on  $N$  as  $q_{\max} \sim N^{1/(\gamma-1)}$  for all values of  $\gamma$ . LPA networks are further characterized by the presence of degree correlations [32]: The average degree of the nearest neighbors of nodes of degree  $q$ ,  $\bar{q}_{nn}(q)$  [33] is of the form  $\bar{q}_{nn}(q) \sim q^{-3+\gamma}$  for  $\gamma < 3$ , and  $\bar{q}_{nn}(q) \sim \ln q$  for  $\gamma > 3$  [31]. See Appendix C for a practical implementation of this model.

By their very construction LPA networks lack a  $K$ -core structure, since the iterative procedure to determine  $K$ -shells for  $K > m$  removes all nodes by exactly reversing the growth process. Therefore, in LPA networks, all nodes belong to the same  $K = m$  shell, where  $m$  is the minimum degree in the network. We thus have  $\langle q \rangle_{K_M} = \langle q \rangle = 2m \ll \sqrt{q_{\max}}$  even for modest values of  $N$ , and according to our generalized prediction, the LEV should scale as  $\sqrt{q_{\max}}$  for all values of  $\gamma$ , in stark opposition to the original CLV scaling, that still predicts in Eq. (3) a different scaling for  $\gamma < 5/2$  and  $\gamma > 5/2$ .

In Fig. 3a we plot the largest eigenvalue obtained in LPA networks with different size and degree exponent  $\gamma$ , as a function of  $\sqrt{q_{\max}}$ . As we can observe, after a short preasymptotic regime for small network sizes (small  $q_{\max}$ ),  $\Lambda_M$  grows as  $\sqrt{q_{\max}}$  for all values of  $\gamma$ , independently of the factor  $\langle q^2 \rangle / \langle q \rangle$ . Interestingly, for all values of  $\gamma$ , the LEV falls onto the same universal curve asymptotically approaching  $\sqrt{q_{\max}}$ , which indicates that this functional form is moreover independent of degree correlations, which change continuously with  $\gamma$

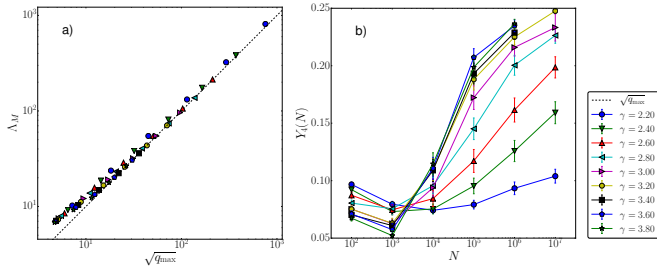


Figure 4. (a) Largest eigenvalue  $\Lambda_M$  as a function of  $\sqrt{q_{\max}}$  in rewired LPA networks with different degree exponent  $\gamma$ . (b) Inverse participation ratio  $Y_4(N)$  as a function of  $N$  in rewired LPA networks with different degree exponent  $\gamma$ . Each point in both graphs corresponds to an average over 100 independent network realization. Error bars are smaller than symbol sizes.

in LPA networks [31]. We conclude that, in perfect agreement with our theory, the spectral properties of LPA networks are ruled alone by the hub. This implies additionally that the PEV is localized around the hub. This fact is verified in Fig. 3b, where we plot the inverse participation ratio  $Y_4(N)$  as a function of  $N$ . In Fig. 3b it turns out clearly that  $Y_4(N)$  goes to a constant for  $N \rightarrow \infty$ , for any degree exponent  $\gamma$  and for sufficiently large  $N$ , indicating that PEV always becomes localized on a set of nodes of finite size (not increasing with  $N$ ): The hub and its immediate neighbors. Further evidence about the localization is provided in Appendix G.

In LPA networks, the lack of a  $K$ -core structure is a fragile property, since it is sufficient to reshuffle connections while preserving the degree of each node [34] to induce some  $K$ -core structure (see Appendix H). This emerging  $K$ -core structure is however not able to restore the scaling predicted by Eq. (2) (see Appendix H). As Fig. 4a shows, reshuffling does not alter the overall behavior, apart from minimal changes: the LEV still scales asymptotically as  $\sqrt{q_{\max}}$  for any  $\gamma$ , while the PEV is still asymptotically localized around the hub, as the inverse participation ratio tending to a constant for large network sizes shows, Fig. 4b.

## V. CONSEQUENCES FOR DYNAMICS ON NETWORKS

The spectral properties of the adjacency matrix determine the behavior of many dynamical processes mediated by topologically complex contact patterns [8–10, 35, 36]. Here we show the consequences that the topological properties uncovered above have for two highly relevant types of dynamics.

### A. Epidemic spreading

The Susceptible-Infected-Susceptible (SIS) model is one of the simplest and most fundamental models for epidemic spreading [37] (see Appendix D for details), showing an epidemic threshold  $\lambda_c$  separating a regime where epidemics get rapidly extinct from a regime where they affect a finite fraction

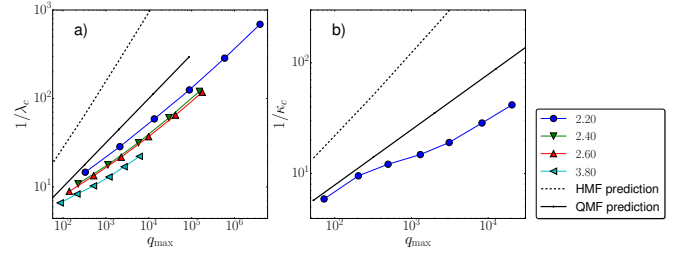


Figure 5. (a) Numerical estimate of the inverse epidemic threshold  $1/\lambda_c$  in LPA as a function of  $q_{\max}$ , for various values of the exponent  $\gamma$ . We consider networks of sizes from  $N = 10^2$  to  $N = 10^8$ . (b) Numerical estimate of the synchronization threshold  $\kappa_c$  for  $\gamma = 2.2$  as a function of the different theoretical predictions. System size ranges from  $N = 300$  to  $N = 300000$ . In both plots the dependence for  $\gamma = 2.2$  predicted by the QMF theory is represented by a thick dashed straight line, and the HMF prediction is represented by a dash-dotted line

of the system. The dependence of this threshold on the network topology is well approximated by the so-called Quenched Mean-Field theory (QMF) (see Appendix D), predicting it to be equal to the inverse of the LEV

$$\lambda_c = \frac{1}{\Lambda_M}. \quad (6)$$

Inserting into this expression the LEV scaling form given by Eq. (2) in the case of random uncorrelated static networks, we see that the threshold always vanishes on power-law distributed networks in the thermodynamic limit, with different scalings depending on the value of  $\gamma$  [35]. For  $\gamma < 5/2$  the expression coincides with the one predicted by the Heterogeneous Mean-Field (HMF) theory [38] (see Appendix D), while HMF theory is violated for  $\gamma > 5/2$ . In LPA networks  $\Lambda_M$  is, for any  $\gamma$ , given by  $\sqrt{q_{\max}}$ , so that Eq. (6) predicts a vanishing of the epidemic threshold qualitatively different from the one on uncorrelated networks for  $\gamma < 5/2$ . In particular, the approach to zero in the thermodynamic limit should be *slower* in LPA networks than in static uncorrelated networks with the same  $\gamma$ .

In order to check this picture, we perform numerical simulations of the SIS model on LPA networks of different degree exponent  $\gamma$ , and determine the threshold using the lifespan method (see Appendix D). In Fig. 5a we plot the numerically estimated threshold as a function of  $q_{\max}$ . We find that the theoretical expectation is followed only approximately: the slopes are smaller than 1 in all cases, the more so for larger values of  $\gamma$ . However, this discrepancy is a finite size effect: as the system size grows the effective slope grows. Asymptotically for large  $N$  the threshold always vanishes as  $\sqrt{q_{\max}}$ , at variance with what happens for uncorrelated static networks for  $\gamma < 5/2$ . The comparison with the slope predicted by HMF theory for  $\gamma = 2.2$  (dashed line) clearly shows the failure of the latter. Hence the remarkable conclusion that on LPA networks the epidemic threshold vanishes asymptotically for any  $\gamma$ , but it never vanishes as predicted by HMF theory, at odds with what happens on static uncorrelated networks.

In the case of real-world networks, our proposed estimate for the scaling of the largest eigenvalue again provides a much



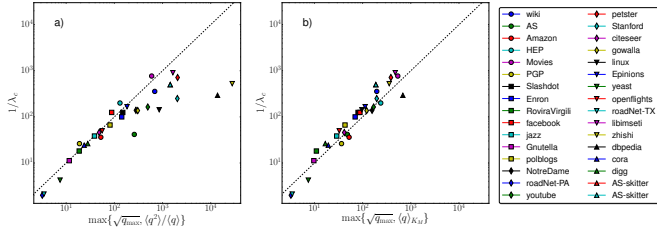


Figure 6. (a) Numerical estimate of the inverse epidemic threshold  $1/\lambda_c$  in real-world networks as a function of the inverse largest eigenvalue approximation  $\max\{\sqrt{q_{\max}}, \langle q^2 \rangle / \langle q \rangle\}$ . (b) Numerical estimate of the inverse epidemic threshold  $1/\lambda_c$  in real-world networks as a function of the inverse improved largest eigenvalue approximation  $\max\{\sqrt{q_{\max}}, \langle q \rangle_{K_M}\}$ . Data for the networks zhishi, NotreDame, roadNet-TX and roadNet-PA are lower bounds to the real threshold, due to computing time limitations.

better overall prediction for the threshold in the SIS model, see Fig. 6, where we compare it with the original CLV prediction. As we can see, in cases where the CLV prediction is off by orders of magnitude, our improved prediction form leads to a much better threshold prediction. As an estimation of the overall goodness of the prediction, the Pearson coefficient for the CLV estimation is barely  $r = 0.344$ , while our improved version yields a much better value  $r = 0.822$ .

## B. Synchronization

Kuramoto dynamics [39] (see Appendix E) is the paradigmatic model for the study of synchronization among weakly coupled oscillators, with application ranging from neural networks to charge-density waves. Its behavior in networks has been investigated in great detail [40, 41], showing the existence of a synchronization threshold for a coupling parameter,  $\kappa_c$ , separating a random from a synchronized phase. Concerning the synchronization threshold, the standard approach is the one in Ref. [8], predicting a synchronized state to appear when the coupling  $\kappa$  among oscillators is larger than the critical value

$$\kappa_c = \frac{k_0}{\Lambda_M}, \quad (7)$$

where  $k_0 = 2/[\pi g(0)]$  and  $g(\omega)$  is the frequency distribution of individual oscillators (see Appendix E). To assess whether the generalized scaling just uncovered for  $\Lambda_M$  on LPA networks has effects also for these dynamics, we perform simulations of the Kuramoto model on growing networks and determine the critical coupling  $\kappa_c$  (see Appendix E for details).

Fig. 5b clearly shows that also for these dynamics the prediction given by the inverse of the LEV is qualitatively correct and, as a consequence, for  $\gamma < 5/2$  the threshold vanishes more slowly than what is predicted for random uncorrelated networks. We conclude that also in this case the nature of the growing network, and in particular the lack of a  $K$ -core structure, has profound consequences for the dynamics mediated by the contact network.

## VI. DISCUSSION

The generalized CLV conjecture we have exposed allows to fully clarify the physical origin of the properties of the adjacency matrix largest eigenpair in complex networks. There are only two subgraphs which determine the LEV and the PEV in a large complex network: the hub with its spokes and the densely mutually interconnected set of nodes singled out as the maximum  $K$ -core [42]. Each of these two subgraphs has (in isolation) an associated LEV: the hub is the center of a star graph and therefore  $\Lambda_M^{(h)} = \sqrt{q_{\max}}$ ; the maximum  $K$ -core is a homogeneous graph and therefore  $\Lambda_M^{(K_M)} = \langle q \rangle_{K_M}$ . The LEV of the global topology is simply given by the largest of the two. In uncorrelated static networks the growth with  $N$  of the two individual LEVs depends on  $\gamma$  and this gives rise to the change of behavior occurring for  $\gamma = 5/2$ , Eq. (3). In growing LPA networks the  $K$ -core structure is by construction absent: The spectral properties are dictated only by the hub (and this remains true also after reshuffling). In generic networks of any origin (and any correlation level), the relation between the average degree of the max  $K$ -core and  $\langle q^2 \rangle / \langle q \rangle$  may break down. However it is still true that the LEV value is the largest between  $\Lambda_M^{(h)}$  and  $\Lambda_M^{(K_M)}$ .

This conjecture is clearly not a proof. However, the understanding of its conceptual origin allows to predict that it should hold for practically all real-world networks. Two possible mechanisms may in principle lead to its breakdown. Either there should be in the network a third, different, type of subgraph, characterized by a LEV larger than both the others. Or there should be an *inhomogeneous* max  $K$ -core in the network, so that  $\Lambda_M^{(K_M)}$  is very different from  $\langle q \rangle_{K_M}$ . Both mechanisms are in principle possible; however they appear to be very unlikely in real-world networks.

Our findings about spectral properties have immediate implications in several contexts. We have shown that properties of dynamical processes as general as epidemics and synchronization are deeply affected by which subgraph determines the LEV. For example, another effect that can be immediately predicted, is that removing the hub may completely disrupt the dynamics when the LEV is given by  $\Lambda_M^{(h)}$  while being practically inconsequential in the other case. Similar consequences are expected to occur for other types of dynamics [9–11, 36]. Another context where these results may have implications is for centrality measures, many of which are variations of the eigenvalue centrality [6, 43]. Finally, it is worth to remark that the example of linear preferential attachment networks clearly points out that the way a network is built may have deep and unexpected implications for its structure and its functionality.

## ACKNOWLEDGMENTS

We acknowledge financial support from the Spanish MINECO, under projects FIS2013-47282-C2-2 and FIS2016-76830-C2-1-P. R.P.-S. acknowledges additional financial support from EC FET-Proactive Project MULTIPLEX (Grant No. 317532) and ICREA Academia, funded by the Generalitat de

Catalunya.

## APPENDICES

### Appendix A: Eigenvector localization and the inverse participation ratio

The concept of the localization of the principal eigenvector  $\{f_i\}$  translates in determining whether the value of its normalized components, satisfying  $\sum_i f_i^2 = 1$ , is evenly distributed among all nodes in the network, that is, if  $f_i \sim N^{-1/2}, \forall i$ , or either it attains a large value on some subset of nodes  $V$ , and is much smaller in all the rest, i.e., if  $f_i \sim N_V^{-1/2}$ , for  $i \in V$  and  $f_i \sim 0$ , for  $i \notin V$ , where  $N_V$  is the size of the localization subset  $V$ .

The presence of localization in the PEV can be easily assessed in ensembles of networks of variable size  $N$  by studying the inverse participation ratio (IPR), defined as [28, 29]

$$Y_4(N) = \sum_{i=1}^N [f_i]^4, \quad (\text{A1})$$

as a function of  $N$ , and fitting its behavior to a power-law decay of the form [25]

$$Y_4(N) \sim N^{-\alpha}. \quad (\text{A2})$$

If the PEV is delocalized, with  $f_i \sim N^{-1/2}, \forall i$ , the exponent  $\alpha$  is equal to 1. Any exponent  $\alpha < 1$  indicates the presence of some form of eigenvector localization, taking place in a sub-extensive set of nodes, of size  $N_V \sim N^\alpha$ . In the extreme case of localization on a single node, or a set of nodes with fixed size, we have  $\alpha = 0$  and  $Y_4(N) \sim \text{const.}$

### Appendix B: The $K$ -core decomposition

The  $K$ -core decomposition [26] is an iterative procedure to classify vertices of a network in layers of increasing density of mutual connections. Starting with the whole graph one removes the vertices with only one connection (degree  $q = 1$ ). This procedure is then repeated until only nodes with degree  $q \geq 2$  are left. The removed nodes constitute the  $K=1$ -shell and those remaining are the  $K=2$ -core. At the next step all vertices with degree  $q = 2$  are removed, thus leaving the  $K=3$ -core. The procedure is repeated iteratively. The maximum  $K$ -core (of index  $K_M$ ) is the set of vertices such that one more iteration of the procedure removes all of them. Notice that all vertices of the  $K$ -core of index  $K$  have degree larger than or equal to  $K$ .

### Appendix C: Building linear preferential attachment networks

Given the mapping of LPA networks with the Price model [44], LPA networks can be easily constructed with the following simplified algorithm [5]: Every time step a new node is

added, with  $m$  new edges. Each one of them is connected to an old node, chosen uniformly at random, with probability  $\phi = a/(a + m)$ ; otherwise, with the complementary probability  $1 - \phi = m/(a + m)$ , the edge is connected to a node chosen with probability proportional to its in-degree  $q_s(t) - m$ . In our simulations we consider LPA networks with minimum degree  $m = 2$  and varying  $\gamma$ , for network sizes ranging from  $N = 10^2$  up to  $N = 10^8$ . Topological and spectral properties of LPA networks are computed averaging over 100 different network configurations for each value of  $\gamma$  and  $N$ .

### Appendix D: Susceptible-Infected-Susceptible epidemic dynamics

The Susceptible-Infected-Susceptible (SIS) model is the simplest model designed to capture the properties of diseases that do not confer immunity [45]. In the SIS model, individuals can be in either of two states, susceptible or infected. Susceptible individuals become infected through a contact with an infected individual at rate  $\beta$ , while infected individuals heal spontaneously at rate  $\mu$ . As a function of the parameter  $\lambda = \beta/\mu$ , the model shows a non-equilibrium phase transition between an active, infected phase for  $\lambda > \lambda_c$ , and an inactive, healthy phase for  $\lambda \leq \lambda_c$ . Interest is placed on the location of the so-called epidemic threshold  $\lambda_c$ , and on its dependence on the topological properties of the network under consideration [46].

Early theoretical approaches to the SIS dynamics were based on the so-called Heterogeneous Mean-Field (HMF) theory [38, 47], which neglects both dynamical and topological correlations by replacing the actual structure of the network, as given by the adjacency matrix, by an annealed version in which edges are constantly rewired, while preserving the degree distribution  $P(q)$ . Within this annealed network approximation [48], a threshold for uncorrelated networks is obtained of the form  $\lambda_c = \langle q \rangle / \langle q^2 \rangle$ . An improvement over this approximate theory is given by Quenched Mean-Field (QMF) theory [7], which, while still neglecting dynamical correlations, takes into account the full structure of the adjacency matrix. Within this approximation, the threshold is given by the inverse of the largest eigenvalue of the adjacency matrix,  $\lambda_c = 1/\Lambda_M$ . Recent and intense activity, based on more sophisticated approaches [35, 49, 50] has shown that on uncorrelated static networks this result is essentially asymptotically correct.

In order to determine  $\lambda_c$  numerically, we resort to the lifespan method [50, 51], which is not affected by the drawbacks that make the consideration of susceptibility unwieldy [49]. Simulations start with only the hub infected. For each run one keeps track of the coverage, i.e. fraction of different nodes which have been touched at least once by the infection. In an infinite network this quantity is vanishing for  $\lambda \leq \lambda_c$ , while it tends asymptotically to 1 in the active region of the phase diagram. In finite networks one can set a threshold  $c$  (we choose  $c = 0.5$ ) and consider all runs that reach a coverage larger than  $c$  as endemic. The average lifespan  $\langle T \rangle$  restricted only to nonendemic runs plays the role of a susceptibility (see Appendix I): The position of the threshold is estimated as the value of  $\lambda$  for which  $\langle T \rangle$  reaches a peak.

## Appendix E: Kuramoto synchronization dynamics

The Kuramoto model [39, 41] describes the dynamics of a collection of weakly coupled nearly identical oscillators. If they are placed on the nodes of a network with adjacency matrix  $A_{ij}$  the equation of motion reads

$$\dot{\theta}_i = \omega_i + \kappa \sum_j A_{ij} \sin(\theta_j - \theta_i), \quad (\text{E1})$$

where  $\kappa$  is a coupling constant and  $\omega_i$  is a quenched random variable (natural frequency), whose distribution  $g(\omega)$  is taken here uniform between  $-1$  and  $1$ . In the initial condition the phases  $\theta_i$  are uniformly random between  $0$  and  $2\pi$ . Defining the global order parameter as

$$r = \left| \frac{1}{N} \sum_i e^{-I\theta_i} \right|, \quad (\text{E2})$$

where  $I$  is the imaginary unit, one finds that there is a critical threshold  $\kappa_c$  separating a disordered phase where  $r = 0$  (in the thermodynamic limit) from a synchronized phase with  $r > 0$ . A QMF-like theory for the Kuramoto model [8] predicts a critical point  $\kappa_c = k_0/\Lambda_M$ , where  $k_0 = 2/[\pi g(0)] = 4/\pi$ , the last equality holding because of the uniform distribution of natural frequencies  $g(\omega)$ .

The value of the critical threshold is numerically determined in finite networks by computing the susceptibility,  $\chi_K(\kappa) = N(\langle r^2 \rangle - \langle r \rangle^2)$ , which shows a peak for  $\kappa = \kappa_c$ .

## Appendix F: Real networks

We consider in our analysis the following real networks datasets:

1. **wiki**: Network of conflicts between editors of the English Wikipedia [52].
2. **AS**: Internet map at the Autonomous System level, collected at the Oregon route server. Vertices represent autonomous systems (aggregations of Internet routers under the same administrative policy), while edges represent the existence of border gateway protocol (BGP) peer connections between the corresponding autonomous systems [53].
3. **Amazon**: Co-purchasing network from the online store Amazon. Nodes represent products, which are joined by edges if they are frequently purchased together [54].
4. **HEP**: Collaboration network between authors of papers submitted to the High Energy Physics section of the online preprint server arXiv. Each node is a scientist. Two scientists are connected by an edge if they have coauthored a preprint [55].
5. **Movies**: Network of movie actor collaborations obtained from the Internet Movie Database (IMDB). Each vertex represents an actor. Two actors are joined by an edge if they have co-starred at least one movie [16].

6. **PGP**: Social network defined by the users of the pretty-good-privacy (PGP) encryption algorithm for secure information exchange. Vertices represent users of the PGP algorithm. An edge between two vertices indicates that each user has signed the encryption key of the other [56].
7. **Slashdot**: User network of the Slashdot technology news website. Nodes represent users, which can tag each other as friends or foes. An edge represents the presence of a tagging between two users [57].
8. **Enron**: Enron email communication network. Nodes represent email addresses. An edge joins two addresses if they have exchanged at least one email [57].
9. **RoviraVirgili**: Email communication network collected at the University Rovira Virgili [58].
10. **facebook**: New Orleans regional Facebook network [59].
11. **jazz**: Network of collaboration among jazz musicians [60].
12. **Gnutella**: Gnutella peer-to-peer file sharing network. Nodes represent hosts in the Gnutella system. An edge stands for a connection between two Gnutella hosts [55].
13. **polblogs**: Network of blogs on US politics [61].
14. **NotreDame**: Notre Dame web graph. Nodes represent web pages from University of Notre Dame. Edges indicate the presence of a hyperlink pointing from one page to another [62].
15. **roadNet-PA**: Road network of Pennsylvania. Intersections and endpoints are represented by nodes, and the roads connecting these intersections or endpoints are represented by undirected edges [57].
16. **youtube**: Friendship network of the video-sharing site Youtube. Nodes are users and an undirected edge between two nodes indicates a friendship [63].
17. **petster**: Network of friendship between users of the website dogster.com [64].
18. **Stanford**: Stanford University web graph [57].
19. **citeseer**: Citation network extracted from the CiteSeer digital library. Nodes are publications and the directed edges denote citations [65].
20. **gowalla**: Network friendship relations from Gowalla, a former location-based social network where users shared their locations. A node represents a user and an edge indicates that a friendship exists between the user represented by the left node and the user represented by the right node [66].
21. **linux**: Network of Linux source code files, with directed edges denoting that they include each other. The network is based on Linux version 3.16 [67].

22. **Epinions**: Trust network from the online social network Epinions. Nodes are users of Epinions and directed edges represent trust between the users [68].
23. **yeast**: Protein-protein interaction network of yeast *S. cerevisiae* [69].
24. **openflights**: Network of flights between airports of the world. A directed edge represents a flight from one airport to another. Dataset extracted from Openflights.org data [70].
25. **roadNet-TX**: Road network of Texas. Intersections and endpoints are represented by nodes, and the roads connecting these intersections or endpoints are represented by undirected edges [57].
26. **libimseti**: Network of ratings given by users of the Czech dating site Libimseti.cz to other users [71].
27. **zhishi**: Network of 'related to' links between articles of the Chinese online encyclopedia Baidu [72].
28. **dbpedia**: The DBpedia network [73].
29. **cora**: Cora citation network [74].
30. **digg**: Reply network of the social news website Digg. Each node in the network is a user of the website, and each directed edge denotes that a user replied to another user [75].
31. **AS-skitter**: Internet topology graph from traceroutes run daily in 2005 [55].

All networks have been considered as undirected, eliminating multiple connections and self-loops from the original datasets. We consider only the giant component of each network. The main topological features of these networks are summarized in Table I

#### Appendix G: Principal eigenvector localization in linear preferential attachment networks

A direct way to observe PEV localization consists in plotting the square of the components  $f_i^2$  as a function of the node degree  $q_i$ , Fig. 7. As we can see from this figure, for all values of  $\gamma$  the component of the PEV associated to the largest values of  $q$  have a macroscopic weight, indicating localization of the PEV in the hubs. This plot presents evidence of a further difference of LPA networks with respect to random uncorrelated networks. In this case, and for  $\gamma < 5/2$ , it is possible to show that the PEV components approach in static networks the form obtained within the annealed network approximation [76], which is given by [25].

$$f_i^{\text{an}} = \frac{q_i}{[N \langle q^2 \rangle]^{1/2}}. \quad (\text{G1})$$

As we can see in Fig. 7, this linear behavior is not present in the data from LPA networks, even for small  $\gamma$  values.

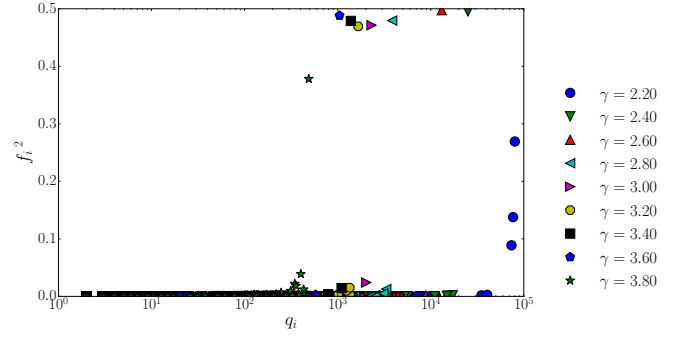


Figure 7. Scatter plot of  $f_i^2$  as a function of the degree  $q_i$  in LPA networks of different degree exponent  $\gamma$ . Network size  $N = 10^6$ .

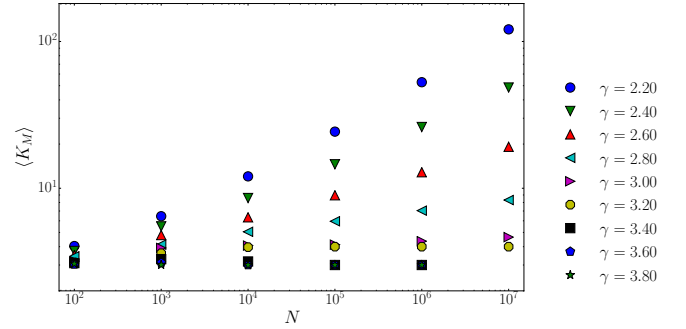


Figure 8. Average maximum core index,  $\langle K_M \rangle$  as a function of network size for reshuffled LPA networks with different degree exponent  $\gamma$ . Error bars are smaller than symbol sizes.

#### Appendix H: $K$ -core structure in reshuffled linear preferential attachment networks

The lack of  $K$ -core structure of LPA networks arises from its peculiar growing nature, in which nodes with minimum degree  $m$  are sequentially attached to the network. This property is not robust, however, since a simple reshuffling procedure can destroy it, inducing a non-trivial  $K$ -core structure. In Fig. 8 we show the average maximum core index,  $\langle K_M \rangle$ , as a function of the network size, computed from LPA networks with different degree exponent, in which edges have been reshuffled according to the degree preserving edge rewiring process described in Ref. [34]. As we can observe, for  $\gamma \geq 3$ , the reshuffling process is not able to induce a substantial  $K$ -core structure. For  $\gamma < 3$ , on the other hand, the  $K$ -core structure generated by rewiring is robust, with an average maximum core index increasing as a power-law with network size [27].

The maximum  $K$ -core resulting from the reshuffling of LPA networks has an average degree,  $\langle q \rangle_{K_M}$ , that depends on the maximum degree  $q_{\text{max}}$ , see Fig. 9. This average degree is however always smaller than  $\sqrt{q_{\text{max}}}$ . Hence the properties of the largest eigenpair are always dictated by the hub, as in the original LPA networks.



Network	$N$	$\langle q \rangle$	$q_{\max}$	$\chi$	$r$	$K_M$	$N_{K_M}$	$\langle q \rangle_{K_M}$	$\chi_{K_M}$
wiki	113123	35.8	20153	18.3	-0.0651	145	936	196.7	0.1049
AS	10790	4.2	2337	61.3	-0.1938	17	34	23.5	0.0513
Amazon	403364	12.1	2752	1.5	-0.0176	10	32886	13.4	0.4025
HEP	11204	21.0	491	5.2	0.6295	238	239	238.0	0.0000
Movies	81860	89.5	3789	5.6	0.2059	359	1125	537.6	0.0721
PGP	10680	4.6	205	3.1	0.2382	31	41	36.5	0.0057
Slashdot	82168	12.3	2552	11.2	-0.0738	55	134	80.5	0.0596
Enron	33696	10.7	1383	12.3	-0.1165	43	275	70.1	0.1584
RoviraVirgili	1133	9.6	71	0.9	0.0782	11	12	11.0	0.0000
facebook	63392	25.8	1098	2.4	0.1768	52	701	88.1	0.1538
jazz	198	27.7	100	0.4	0.0202	29	30	29.0	0.0000
Gnutella	62561	4.7	95	1.5	-0.0927	6	1004	9.1	0.1253
polblogs	1222	27.4	351	2.0	-0.2213	36	55	43.2	0.0139
NotreDame	325729	6.7	10721	40.9	-0.0534	155	1367	157.3	0.0591
roadNet-PA	1087562	2.8	9	0.1	0.1220	3	916	3.3	0.0208
youtube	1134890	5.3	28754	92.9	-0.0369	51	845	86.1	0.2458
petster	426485	40.1	46503	50.3	-0.0884	248	1177	386.6	0.1198
Stanford	255265	15.2	38625	132.5	-0.1156	71	387	112.2	0.3004
citeseer	365154	9.4	1739	4.1	-0.0632	15	1850	25.4	0.3787
gowalla	196591	9.7	14730	30.7	-0.0293	51	185	76.0	0.0957
linux	30817	13.8	9338	60.6	-0.1747	23	439	41.3	0.7557
Epinions	75877	10.7	3044	16.2	-0.0406	67	485	113.0	0.1608
yeast	1458	2.7	56	1.7	-0.2095	5	6	5.0	0.0000
openflights	2905	10.8	242	4.2	0.0489	28	38	32.8	0.0079
roadNet-TX	1351137	2.8	12	0.1	0.1271	3	1491	3.4	0.0495
libimseti	220970	156.0	33389	9.5	-0.1390	273	4110	475.6	0.2285
zhishi	372840	12.4	127066	2243.5	-0.2825	228	449	282.7	0.0582
dbpedia	3915921	6.4	469692	2156.2	-0.0427	20	70	27.9	0.0620
cora	23166	7.7	377	2.1	-0.0553	13	25	17.4	0.0329
digg	29652	5.7	283	3.9	0.0027	9	1339	16.8	0.4619
AS-skitter	1694616	13.1	35455	109.4	-0.0814	111	222	150.0	0.0451

Table I. Topological properties of the real networks considered: Network size  $N$ ; average degree  $\langle q \rangle$ ; maximum degree  $q_{\max}$ ; heterogeneity parameter  $\chi = \langle q^2 \rangle / \langle q \rangle^2 - 1$ ; degree correlations as measured by the Pearson coefficient  $r$  [32]; maximum core index  $K_M$ ; size of the maximum core  $N_{K_M}$ ; average internal degree of the maximum core  $\langle q \rangle_{K_M}$ ; heterogeneity parameter of the maximum core  $\chi_{K_M} = \langle q^2 \rangle_{K_M} / \langle q \rangle_{K_M}^2 - 1$ .

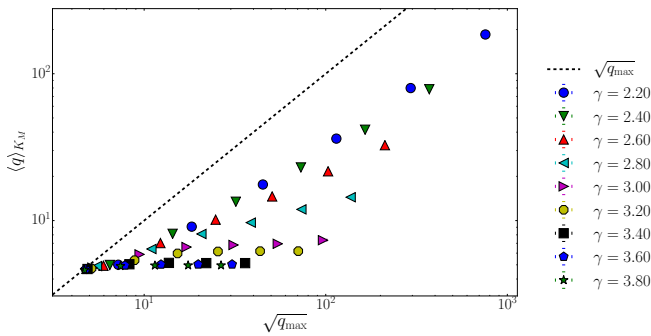


Figure 9. Average maximum core index,  $\langle K_M \rangle$  as a function of network size for reshuffled LPA networks with different degree exponent  $\gamma$ . Error bars are smaller than symbol sizes.

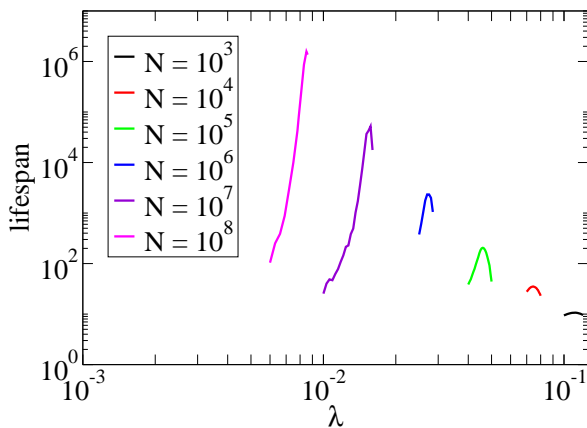


Figure 10. Average lifespan vs spreading rate  $\lambda_c$  of SIS epidemics starting from a single infected node (the hub) and reaching the healthy absorbing state before the coverage reaches the threshold value  $c = 0.5$ . Data are for LPA networks with  $\gamma = 2.6$  and various system size  $N$ .

## Appendix I: Lifespan method for the determination of the SIS epidemic threshold

In Fig. 10 we plot the “susceptibility”  $\langle T \rangle$  that allows to determine the position of the threshold for SIS epidemics on LPA networks.

- 
- [1] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, U.K., 2011).
  - [2] A. N. Samukhin, S. N. Dorogovtsev, and J. F. F. Mendes, *Phys. Rev. E* **77**, 036115 (2008).
  - [3] S. Fortunato, *Physics Reports* **486**, 75 (2010).
  - [4] B. Karrer, M. E. J. Newman, and L. Zdeborová, *Phys. Rev. Lett.* **113**, 208702 (2014).
  - [5] M. Newman, *Networks: An Introduction* (Oxford University Press, Inc., New York, NY, USA, 2010).
  - [6] P. Bonacich, *Journal of Mathematical Sociology* **2**, 113 (1972).
  - [7] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, *ACM Transactions on Information and System Security (TISSEC)* **10**, 1 (2008).
  - [8] J. G. Restrepo, E. Ott, and B. R. Hunt, *Phys. Rev. E* **71**, 036151 (2005).
  - [9] J. G. Restrepo, E. Ott, and B. R. Hunt, *Phys. Rev. Lett.* **100**, 058701 (2008).
  - [10] A. Pomerance, E. Ott, M. Girvan, and W. Losert, *Proceedings of the National Academy of Sciences* **106**, 8209 (2009).
  - [11] O. Kinouchi and M. Copelli, *Nat Phys* **2**, 348 (2006).
  - [12] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. E* **68**, 046109 (2003), arXiv:0306340 [cond-mat].
  - [13] D.-H. Kim and A. E. Motter, *Phys. Rev. Lett.* **98**, 248701 (2007).
  - [14] R. R. Nadakuditi and M. Newman, *Phys. Rev. E* **87**, 012803 (2013).
  - [15] J. G. Restrepo, E. Ott, and B. R. Hunt, *Phys. Rev. E* **76**, 056119 (2007).
  - [16] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
  - [17] F. Chung, L. Lu, and V. Vu, *Proc. Natl. Acad. Sci. USA* **100**, 6313 (2003).
  - [18] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, *Euro. Phys. J. B* **38**, 205 (2004).
  - [19] M. Boguñá and R. Pastor-Satorras, *Phys. Rev. E* **68**, 036112 (2003).

- [20] F. Chung, L. Lu, and V. Vu, *Ann. Comb.* **7**, 21 (2003).
- [21] S. N. Dorogovtsev and J. F. F. Mendes, *Advances in Physics* **51**, 1079 (2002).
- [22] G. H. Golub and C. F. Van Loan, *Matrix computations*, Vol. 3 (JHU Press, 2012).
- [23] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, *Phys. Rev. E* **71**, 027103 (2005).
- [24] C. Castellano and R. Pastor-Satorras, *Scientific Reports* **2**, 00371 (2012).
- [25] R. Pastor-Satorras and C. Castellano, *Sci. Rep.* **6**, 18847 (2016), arXiv:1505.06024.
- [26] S. B. Seidman, *Social Networks* **5**, 269 (1983).
- [27] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Phys. Rev. Lett.* **96**, 040601 (2006).
- [28] A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, and J. F. F. Mendes, *Phys. Rev. Lett.* **109**, 128702 (2012).
- [29] T. Martin, X. Zhang, and M. E. J. Newman, *Phys. Rev. E* **90**, 052808 (2014).
- [30] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Physical Review Letters* **85**, 4633 (2000).
- [31] A. Barrat and R. Pastor-Satorras, *Physical Review E* **71**, 036127 (2005).
- [32] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [33] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).
- [34] S. Maslov and K. Sneppen, *Science* **296**, 910 (2002).
- [35] C. Castellano and R. Pastor-Satorras, *Phys. Rev. Lett.* **105**, 218701 (2010).
- [36] D. B. Larremore, W. L. Shew, and J. G. Restrepo, *Phys. Rev. Lett.* **106**, 058101 (2011).
- [37] R. M. Anderson and R. M. May, *Infectious diseases in humans* (Oxford University Press, Oxford, 1992).
- [38] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
- [39] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, *Rev. Mod. Phys.* **77**, 137 (2005).
- [40] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Physics Reports* **469**, 93 (2008).
- [41] F. A. Rodrigues, T. K. D. Peron, P. Ji, and J. Kurths, *Phys. Rep.* **610**, 1 (2016), arXiv:1511.07139.
- [42] Notice that in many cases the hub actually is part of the maximum  $K$ -core. Nevertheless there is a clear distinction between the case the PEV is localized on it and its immediate neighbors or the PEV is localized around the maximum  $K$ -core as a whole.
- [43] L. Katz, *Psychometrika* **18**, 39 (1953).
- [44] D. J. de Solla Price, *J. Amer. Soc. Inform. Sci.* **27**, 292 (1976).
- [45] M. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2007).
- [46] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Rev. Mod. Phys.* **87**, 925 (2015).
- [47] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. E* **63**, 066117 (2001).
- [48] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Rev. Mod. Phys.* **80**, 1275 (2008).
- [49] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. E* **86**, 041125 (2012).
- [50] M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. Lett.* **111**, 068701 (2013).
- [51] A. S. Mata, M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. E* **91**, 052117 (2015).
- [52] U. Brandes and J. Lerner, *J. Classification* **27**, 279 (2010).
- [53] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A statistical physics approach* (Cambridge University Press, Cambridge, 2004).
- [54] J. Leskovec, L. Adamic, and B. Huberman, *ACM Transactions on the Web* **1**, 5 (2007).
- [55] J. Leskovec, J. Kleinberg, and C. Faloutsos, *ACM Trans. Knowl. Discov. Data* **1**, 2 (2007).
- [56] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, *Phys. Rev. E* **70**, 056122 (2004).
- [57] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney, *Internet Mathematics* **6**, 29 (2009).
- [58] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, *Phys. Rev. E* **68**, 065103 (2003).
- [59] B. Viswanath, A. Mislove, M. Cha, and K. h. P. Gummadi, in *Proceedings of the 2nd ACM workshop on Online social networks* (ACM, 2009) pp. 37–42.
- [60] P. M. Gleiser and L. Danon, *Advances in Complex Systems* **6**, 565 (2003).
- [61] L. A. Adamic and N. Glance, in *Proceedings of the 3rd international workshop on Link discovery* (ACM, 2005) pp. 36–43.
- [62] R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **401**, 130 (1999).
- [63] J. Yang and J. Leskovec, *Knowledge and Information Systems* **42**, 181 (2015).
- [64] “Dogster friendships network dataset – KONECT,” (2016).
- [65] K. D. Bollacker, S. Lawrence, and C. L. Giles, in *Proceedings of the Second International Conference on Autonomous Agents, AGENTS ’98* (ACM, New York, NY, USA, 1998) pp. 116–123.
- [66] E. Cho, S. A. Myers, and J. Leskovec, in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’11* (ACM, New York, NY, USA, 2011) pp. 1082–1090.
- [67] “Linux network dataset – KONECT,” (2016).
- [68] M. Richardson, R. Agrawal, and P. Domingos, “Trust management for the semantic web,” in *The Semantic Web - ISWC 2003: Second International Semantic Web Conference, Sanibel Island, FL, USA, October 20-23, 2003. Proceedings*, edited by D. Fensel, K. Sycara, and J. Mylopoulos (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003) pp. 351–368.
- [69] H. Jeong, S. P. Mason, A. L. Barabasi, and Z. N. Oltvai, *Nature* **411**, 41 (2001).
- [70] T. Opsahl, F. Agneessens, and J. Skvoretz, *Social Networks* **32**, 245 (2010).
- [71] J. Kunegis, G. Gröner, and T. Gottron, in *Proceedings of the 4th ACM RecSys Workshop on Recommender Systems and the Social Web, RSWeb ’12* (ACM, New York, NY, USA, 2012) pp. 37–44.
- [72] X. Niu, X. Sun, H. Wang, S. Rong, G. Qi, and Y. Yu, in *Proceedings of the 10th International Conference on The Semantic Web - Volume Part II, ISWC’11* (Springer-Verlag, Berlin, Heidelberg, 2011) pp. 205–220.
- [73] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, in *Proceedings of the 6th International The Semantic Web and 2Nd Asian Conference on Asian Semantic Web Conference, ISWC’07/ASWC’07* (Springer-Verlag, Berlin, Heidelberg, 2007) pp. 722–735.
- [74] L. Šubelj and M. Bajec, in *Proceedings of the 22Nd International Conference on World Wide Web, WWW ’13 Companion* (ACM, New York, NY, USA, 2013) pp. 527–530.
- [75] M. De Choudhury, H. Sundaram, A. John, and D. D. Seligmann, in *Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04, CSE ’09* (IEEE Computer Society, Washington, DC, USA, 2009) pp. 151–158.
- [76] M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. E* **79**, 036110 (2009).